CLAIMS

What is claimed is:

1. A method for synthesizing an auditory scene, comprising the steps of:

(a) dividing an input audio signal into a plurality of different frequency bands; and

(b) applying two or more different sets of one or more spatial parameters to two or more of the different frequency bands in the input audio signal to generate two or more synthesized audio signals of the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set of one or more spatial parameters is applied to the input audio signal as if the input audio signal corresponded to a single audio source in the auditory scene.

2. The invention of claim 1, wherein each set of one or more spatial parameters corresponds to a different audio source in the auditory scene.

3. The invention of claim 1, wherein, for at least one of the sets of one or more spatial parameters, at least one of the spatial parameters corresponds to a combination of two or more different audio sources in the auditory scene that takes into account relative dominance of the two or more different audio sources in the auditory scene.

4. The invention of claim 1, wherein the input audio signal is a mono signal.

5. The invention of claim 4, wherein the mono signal corresponds to a combination of two or more different mono source signals, wherein the two or more different frequency bands are selected by comparing magnitudes of the two or more different mono source signals, wherein, for each of the two or more different frequency bands, one of the mono source signals dominates the other mono source signals.

6. The invention of claim 4, wherein the mono signal corresponds to a combination of left and right audio signals of a binaural signal, wherein each different set of one or more spatial parameters is generated by comparing the left and right audio signals in a corresponding frequency band.

7. The invention of claim 1, wherein step (a) comprises the step of dividing the input audio signal into the plurality of different frequency bands based on information corresponding to the different sets of one or more spatial parameters.

8. The invention of claim 1, wherein each set of one or more spatial parameters is applied to at least one frequency band in which the input audio signal is dominated by a corresponding audio source in the auditory scene.

9. The invention of claim 1, wherein each set of one or more spatial parameters comprises one or more of an interaural level difference, an interaural time delay, and a head-related transfer function.

10. The invention of claim 1, wherein:

step (a) further comprises the step of converting the input audio signal from a time domain into a frequency domain; and

step (b) further comprises the step of converting the two or more synthesized audio signals from the frequency domain into the time domain.

11. The invention of claim 1, wherein the two or more synthesized audio signals comprise left and right audio signals of a binaural signal corresponding to the auditory scene.

12. The invention of claim 1, wherein the two or more synthesized audio signal comprise two or more signals of a multi-channel audio signal corresponding to the auditory scene.

13. The invention of claim 1, wherein:

the input audio signal is a mono signal;

each set of one or more spatial parameters corresponds to a different audio source in the auditory scene;

step (a) comprises the steps of:

(1) converting the mono signal from a time domain into a frequency domain;

(2) dividing the converted mono signal into the plurality of different frequency bands based on information corresponding to the sets of one or more spatial parameters;

each set of one or more spatial parameters is applied to at least one frequency band in which the input audio signal is dominated by a corresponding audio source in the auditory scene;

each set of one or more spatial parameters comprises one or more of an interaural level difference, an interaural time delay, and a head-related transfer function;

the two or more synthesized audio signals comprise left and right audio signals of a binaural signal corresponding to the auditory scene; and

step (b) further comprises the step of converting the left and right audio signals from the frequency domain into the time domain.

14. The invention of claim 13, wherein the mono signal corresponds to a combination of two or more different mono source signals, wherein the two or more different frequency bands are selected by comparing magnitudes of the two or more different mono source signals, wherein, for each of the two or more different frequency bands, one of the mono source signals dominates the other mono source signals.

15. The invention of claim 13, wherein the mono signal corresponds to a combination of left and right audio signals of a binaural signal, wherein each different set of one or more spatial parameters is generated by comparing the left and right audio signals in a corresponding frequency band.

16. A machine-readable medium, having encoded thereon program code, wherein, when the program code is executed by a machine, the machine implements a method for synthesizing an auditory scene, comprising the steps of:

(a) dividing an input audio signal into a plurality of different frequency bands; and

(b) applying two or more different sets of one or more spatial parameters to two or more of the different frequency bands in the input audio signal to generate two or more synthesized audio signals of the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set of one or more spatial parameters is applied to the input audio signal as if the input audio signal corresponded to a single audio source in the auditory scene.

17. An apparatus for synthesizing an auditory scene, comprising:

(a) means for dividing an input audio signal into a plurality of different frequency bands; and

(b) means for applying two or more different sets of one or more spatial parameters to two or more of the different frequency bands in the input audio signal to generate two or more synthesized audio signals of the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set of one or more spatial parameters is applied to the input audio signal as if the input audio signal corresponded to a single audio source in the auditory scene.

18. An apparatus for synthesizing an auditory scene, comprising:

(1) an auditory scene synthesizer configured to:

(a) divide an input audio signal into a plurality of different frequency bands; and

(b) apply two or more different sets of one or more spatial parameters to two or more of the different frequency bands in the input audio signal to generate two or more synthesized audio signals of the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set of one or more spatial parameters is applied to the input audio signal as if the input audio signal corresponded to a single audio source in the auditory scene; and

(2) one or more inverse time-frequency transformers configured to convert the two or more synthesized audio signals from a frequency domain into a time domain.

19. A method for processing two or more input audio signals, comprising the steps of:

(a) converting the two or more input audio signals from a time domain into a frequency domain;

(b) generating a set of one or more auditory scene parameters for each of two or more different frequency bands in the two or more converted input audio signals, where each set of one or more auditory scene parameters is generated as if the corresponding frequency band corresponded to a single audio source in an auditory scene; and

(c) combining the two or more input audio signals to generate a combined audio signal.

20. The invention of claim 19, wherein:

the two or more input audio signals are mono signals corresponding to different audio sources in the auditory scene;

each set of one or more auditory scene parameters corresponds to an audio source that dominates the other audio sources in the corresponding frequency band; and

the two or more input audio signals are combined in the time domain to generate the combined audio signal.

21. The invention of claim 19, wherein:

the two or more input audio signals are left and right audio signals of a binaural signal;

each set of one or more auditory scene parameters is generated by comparing the left and right audio signals in the corresponding frequency band;

the combined audio signal is generated by performing auditory scene removal on the left and right audio signals in the frequency domain based on the two or more sets of one or more auditory scene parameters; and

further comprising the step of converting the combined audio signal from the frequency domain into the time domain.

22. A machine-readable medium, having encoded thereon program code, wherein, when the program code is executed by a machine, the machine implements a method for processing two or more input audio signals, comprising the steps of:

(a) converting the two or more input audio signals from a time domain into a frequency domain;

(b) generating a set of one or more auditory scene parameters for each of two or more different frequency bands in the two or more converted input audio signals, where each set of one or more auditory scene parameters is generated as if the corresponding frequency band corresponded to a single audio source in an auditory scene; and

(c) combining the two or more input audio signals to generate a combined audio signal.

23. An apparatus for processing two or more input audio signals, comprising:

(a) means for converting the two or more input audio signals from a time domain into a frequency domain;

(b) means for generating a set of one or more auditory scene parameters for each of two or more different frequency bands in the two or more converted input audio signals, where each set of one or more auditory scene parameters is generated as if the corresponding frequency band corresponded to a single audio source in an auditory scene; and

(c) means for combining the two or more input audio signals to generate a combined audio signal.

24. An apparatus for processing two or more input audio signals, comprising:

(a) a time-frequency transformer configured to convert the two or more input audio signals from a time domain into a frequency domain;

(b) an auditory scene parameter generator configure to generate a set of one or more auditory scene parameters for each of two or more different frequency bands in the two or more converted input audio signals, where each set of one or more auditory scene parameters is generated as if the corresponding frequency band corresponded to a single audio source; and

(c) a combiner configured to combine the two or more input audio signals to generate a combined audio signal.

25. The invention of claim 24, wherein:

the two or more input audio signals are mono signals corresponding to different audio sources in the auditory scene;

each set of one or more auditory scene parameters corresponds to an audio source that dominates the other audio sources in the corresponding frequency band; and

the combiner operates in the time domain.

26. The invention of claim 24, wherein:

the two or more input audio signals are left and right audio signals of a binaural signal;

each set of one or more auditory scene parameters is generated by comparing the left and right audio signals in the corresponding frequency band;

the combiner is configured to perform auditory scene removal on the left and right audio signals in the frequency domain based on the two or more sets of one or more auditory scene parameters; and

further comprising an inverse time-frequency transformer configured to convert the combined audio signal from the frequency domain into the time domain.